DOCUMENT RESUME

ED 302 071 FL 017 694

AUTHOR Stromswold, Karin
TITLE Linguistic Representations of Children's Wh-Questions.
PUB DATE Aug 88
NOTE 9p.; In: Papers and Reports on Child Language Development, Volume 27; see FL 017 572.
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Child Language; *Language Acquisition; Language Research; Pragmatics; Preschool Children; *Sentence Structure; Verbal Learning
IDENTIFIERS *Wh Questions

ABSTRACT

A study examined 12 preschool children's early use of "who," "what," and "which" questions in spontaneous speech. Results indicated that children began to ask object questions before they asked subject questions, and acquired argument questions before adjunct questions. It was suggested that the two results could reflect the syntactic structure of the questions or the learning mechanism of the children, but that the results were most consistent with the hypothesis that children are able to theta-govern traces before they are able to antecedent-govern traces. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED302071

PRCLD 27 (1988)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
E. Clark
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Linguistic Representations of Children's *Wh*-Questions[1]

Karin Stromswold

Massachusetts Institute of Technology and Harvard Medical School

Abstract. I determined when 12 children acquired different types of questions. I found that they began to ask object questions before they began to ask subject questions. I also found that they acquired argument questions before adjunct questions. These two results could be the reflection of the syntactic structure of the questions or the learning mechanism of the children. It is possible that the children mastered object questions before subject or adjunct questions because object traces are theta-governed whereas adjunct and subject traces are antecedent-governed. Another possible explanation is that they acquired object questions first because object gaps are more obvious than subject or adjunct gaps. While the data are not conclusive, they are more consistent with the former explanation.

1. Do children's subject questions have gaps?

Introduction

A subject question asks a question about the identity of the subject of a sentence while an object question asks a question about the object of a sentence. One way subject and object questions differ is that an object question clearly has a gap while a subject question may or may not have a gap. Consider for a moment questions that have the verb *meet*. The verb *meet* requires a subject and an object. Therefore, the only viable structure for the object question *who will Mary meet* is one in which there is a gap following the verb. This gap is somehow linked to the sentence-initial *who* (*who*$_i$ *will Mary meet* t_i*.?*).

Now compare the subject question *who will meet John?* with the sentence *Mary will meet John*. There are two possible structures for this question. One possibility is that there is no gap and the *who* is in the same place as *Mary* (*who will meet John?*). Alternatively, there could be an invisible gap between *who* and *will* (*who*$_i$ t_i *will meet John?*). Which structure is correct? According to Government-Binding (GB) theory (Chomsky, 1986), both subject and object questions have gaps. According to Generalized Phrase Structure Grammar (Gazdar, 1981; Gazdar *et. al.*, 1985), subject questions do not have gaps.

Gazdar (1981) predicts that, because subject questions are gapless, they will be acquired before object questions. He writes (1981, p. 172):

> A derived category [a gap] has to contain twice as much syntactic information as a basic category. This suggests that structures which involve derived categories will impose a heavier processing load than those that do not. If this is so, then our analysis predicts that matrix subject relatives and questions will be significantly easier to process than all other relatives and questions. This prediction is borne out, both developmentally and for adult speakers by recent psycholinguistic work.

All of the developmental research (e.g., Ervin-Tripp,1970; Cairns and Hsu, 1978; Tyack and Ingram, 1977; and Stewart and Sinclair, 1975) examines children's **comprehension** of subject and object questions. Contrary to what Gazdar suggests, this research does not uniformly show that children comprehend subject questions earlier or

FL 017694

with greater ease than object questions.[2] In addition, comprehension research may not be the most appropriate means to determine whether subject questions have gaps. Research by Wanner and Maratsos (1978) and Frazier, Clifton, and Randall (1983) indicates that the greater the distance between a word and its gap, the more difficulty adults have processing the sentence. Therefore, even if subject questions **do** have gaps, children might comprehend them more readily than object questions because the distance between the *wh*-word and its gap would be less for subject questions than for object questions.

If Gazdar is right and subject questions are gapless, then (all else being equal) subject questions should be acquired earlier and used with greater ease than object questions. In Analysis 1, I determined when 12 children first produced subject and object questions. In Analysis 2, I examined how often they produced subject and object question. In Analysis 3, I determined whether the order of acquisition of subject and object questions was merely a reflection of adult input frequencies. In Analysis 4, I attempted to determine whether the order of acquisition was actually just an artifact of sampling.

Analysis 1. Age of acquisition of subject and object questions

I examined the spontaneous speech transcripts of the 12 children listed in Figure 1 and determined when each child asked her first subject and object *who*, *what*, and *which* questions.[3]

Figure 1: Children included in the transcript analyses

Corpus collected by	Child	Ages	# lines w/ *wh*-words
Brown (1973):	Adam	2;3-5;2	4,859
	Sarah	2;3-5;1	2,350
	Eve	1;6-2;3	651
MacWhinney (n.d):	Ross	2;10-6;6	2,372
	Mark	1;5-4;7	830
Snow (1983):	Nathan	2;6-3;9	1,724
Sachs (1983):	Naomi	1;2-4;9	1,023
Bloom (1973):	Peter	1;10-3;2	1,920
	Allison	1;4-2;10	37
Higginson (1985):	April	1;10-2;11	263
	May	0;11-0;11	0
	June	1;3-1;9	37
Total number of lines with *wh*-words:			16,066

Results. Overall, the children asked their first object question 1.5 months **before** they asked their first subject question. This difference was significant by sign test ($p < .02$) and marginally significant by t-test (t (9 d.f.) = -2.04, $p < .07$).[4] On average, they asked their first *who* subject question 0.5 months before they asked their first *who* object question. This difference was not significant by either sign ($p > .34$) or t-tests ($p > .26$). On average, they asked their first *what* object question 2.5 months before they asked their first *what* subject question. This difference was marginally significant by sign test ($p < .06$) and t-test (t (9 d.f.) = -2.04, $p < .07$). Lastly, they asked their first *which* object question a mean of 8.8 months before they asked their first *which* subject question. This

difference was marginally significant by sign test ($p < .06$) and significant by t-test ($t(3$ d.f.$)= -3.51$, $p < .04$).

Analysis 2. Relative frequency of subject and object questions

I determined the frequency of subject and object *who*, *what*, and *which* questions in the transcribed speech of each of the 12 children shown in Figure 1.

Results. Overall, the children produced 3 times as many object questions as subject questions. Thirty-three percent of their *who* questions , 88% of their *what* questions, and 77% of their *which* questions were object questions.

Analysis 3. Adult input of subject and object questions

Perhaps the children acquired object questions before subject questions because the people who spoke to them asked more object questions than subject questions. For each of the 12 children, I determined how frequently the adults speaking to each child asked subject and object *who* and *what* questions. I determined the percent of adult *who* and *what* questions that were subject questions [5]

Results. I found no significant correlation between the percent of adult and child *who* questions which were subject questions ($r=-.09$). I also found no significant correlation between the percent of adult and child *what* questions which were subject questions ($r=.26$). In addition, the adult percentages did not correlate significantly with the ages at which the children acquired subject or object *who* or *what* questions.(all correlation coefficients were between -.52 and +.39).

Analysis 4. Sampling artifacts and subject and object questions

Perhaps the children only appeared to acquire object questions before subject questions because of sampling. If they asked many more object questions than subject questions, then a random sample of their speech would be more likely to include object questions than subject questions. If this is why object questions appeared first in the transcripts, then we would expect that the types of questions that appeared frequently in the transcripts would be "acquired" earlier and questions that were infrequent would be "acquired" late. I determined whether the children's frequencies and ages of acquisition were correlated in this manner.

Results. This pattern of results was not obtained. Frequent use of subject *who* questions was fairly correlated with early acquisition of subject *who* questions ($r = -.81$). However, frequent use of object *who* questions correlated with **late** acquisition of *who* object questions ($r = .74$). The same pattern held for *what* questions, though the correlations were not even marginally significant. ($r = -.40$ for subject *what* questions and $r= .31$ for object *what* questions).

Discussion

Gazdar (1981) predicts that subject questions should be easier to learn and use than object questions because subject questions are gapless. However, the production data discussed in Analyses 1 and 2 suggest that children find subject questions **harder** than object questions. Thus, the results of Analyses 1 and 2 suggest that Gazdar is wrong on at least one of two points. Either gaps don't make a structure harder to acquire and use, or subject questions have gaps.

It is somewhat surprising that the children used object questions earlier and more frequently than subject questions. There are at least 5 reasons why we might expect subject questions to be easier than object questions First, the distance between the *wh*-word and

the gap is shorter in subject questions than object questions. Second, subject questions generally have fewer words than object questions. Third, subject questions don't have overt subject-auxiliary inversion. Fourth, subject questions never require *do*-support. Fifth, subject questions are homologous in structure to simple declaratives.

The children might have acquired object questions before subject questions because they heard more object questions than subject questions. The results of Analysis 3 suggest that this is not the correct explanation. Another possible explanation is that object questions only appear to have been acquired before subject questions because of speech sampling. While it is not possible to rule out this explanation, the results of Analysis 4 suggest it is not correct.

2. Do children's argument and adjunct questions differ structurally?

Introduction

Argument questions ask questions about an argument of a sentence. Argument questions include all *who* and *what* questions and certain *where* questions (e.g., *where did Mary put the book?*, *where do the books go?*, *where is the book?*, etc.) and *how* questions (e.g., *how big is the book?*, *how many did she have?*, etc.). Adjunct questions ask questions about an adjunct in a sentence. Adjunct questions include all *why*, *when*, and *how come* questions, and most *where* and *how* questions (e.g., *where did Mary meet John?* and *how did Mary know John?*).

Is there any developmental evidence to suggest that argument and adjunct questions differ structurally? Many researchers (e.g, Brown, 1973; Ervin-Tripp, 1970; Tyack and Ingram, 1977; Labov and Labov, 1978) have reported that children usually begin to ask *who* and *what* questions (i.e., argument questions) before they ask *why*, *when*, and *how* questions (i.e., adjunct questions). However, no one has done a longitudinal study designed to test whether argument questions are acquired before adjunct questions. The following is such a study. In Analysis 1, I compared the average age of acquisition of argument and adjunct questions. Analysis 2 is an attempt to tease apart the conceptual and syntactic factors which may influence when argument and adjunct questions are acquired. In it, I compared when argument and adjunct *where* and *how* questions were acquired. In Analysis 3, I attempted to determine whether children's difficulty with adjunct questions reflects a general difficulty with **adjuncts** or a specific difficulty with adjunct **questions**. I did this by comparing the acquisition of locative questions and declaratives.

Analysis 5. The mean age of acquisition for argument & adjunct questions

I recorded the age at which each of the 12 children began producing nonroutine, sentential examples of each type of question. For each child, I calculated the mean age of acquisition for argument questions and adjunct questions. For argument questions, I did this by averaging together the age of acquisition for *who*, *what*, *which*, argument *where*, and argument *how* questions. For adjunct questions, I averaged together the age of acquisition for *why*, *when*, *how come*, adjunct *where*, and adjunct *how* questions. I then compared the average age of acquisition for argument and adjunct questions.

Results. Averaging across *wh*-words and children, the children asked their first argument question 7.1 months before they asked their first adjunct question. This difference was statistically significant by both the sign ($p < .001$) and t-tests ($t(9) = 6.10$, $p < .0001$).

Analysis 6. Syntactic versus conceptual differences and difficulties

The children might have found adjunct questions more difficult than argument questions for conceptual reasons or structural reasons. One way to determine whether syntactic or conceptual complexity delayed the acquisition of adjunct questions is to compare when they acquired argument and adjunct *where* questions. Argument and adjunct *where* questions are both questions about locations. Adjunct locations aren't any more conceptually complicated than argument locations. Therefore, any difference in age of acquisition between adjunct and argument *where* questions cannot be due to differences in conceptual complexity. The same logic holds for argument and adjunct *how* questions. Any difference in the age of acquisition for argument and adjunct *how* questions cannot be attributed to differences in conceptual complexity. I compared when each of the 12 children acquired argument and adjunct *where* and *how* questions.

Results. On average, the children asked their first argument *where* question 6.8 months before they asked their first adjunct *where* question. This difference was significant by sign test ($p < .002$) and t-test ($t(8)= 3.49$, $p < .008$). On average, they asked their first argument *how* question 7.5 months before their first adjunct *how* question. This difference was also significant by sign test ($p < .031$) and t-test ($t(5) = 3.45$, $p < .018$). The difference for *how* questions is particularly striking for two reasons. First, most of the children's early *how* argument questions had *how* phrases (e.g., *how big is the book?*) while most their early *how* adjunct questions just had *how* (e.g., *how does she swim?*). Second, most of their early argument *how* questions were quantificational queries, something which is conceptually quite sophisticated.

Analysis 7. Acquisition of argument and adjunct questions and declaratives

Perhaps the children acquired adjunct questions late because of a general problem with adjuncts rather than because of a specific problem with adjunct questions. These two possibilities can be teased apart by comparing when they acquired locative arguments and adjuncts in declaratives with when they acquired *where* argument and adjunct questions. If they acquired *where* adjunct questions late because they acquired declarative adjuncts late, then we would expect to find a significant difference in the age of acquisition for argument and adjunct locative declaratives just as we found a significant difference for *where* questions. We would also expect that the children would have acquired adjunct questions soon after they acquired adjunct declaratives. I determined when each of the 12 children acquired locative arguments and adjuncts in declaratives by searching for all of the children's lines which had the prepositions *in*, *on*, or *at*.[6]

Results. The children began using declarative argument locatives an average of 0.5 months before they began using declarative adjunct locatives. This difference was not significant by either sign ($p<.50$) or t-tests ($t(9) = 0.75$, $p > .475$). In contrast, they acquired *where* argument questions 6.8 months before *where* adjunct questions, a difference which was significant by both sign test ($p < .002$) and t-test ($t(8)= 3.49$, $p < .008$). On average, *where* argument questions appeared 3.5 months after declarative argument locatives and *where* adjunct questions appeared 10.7 months after declarative adjunct locatives.

Discussion

The results of Analysis 5 indicate that the children acquired argument questions at a significantly earlier age than they acquired adjunct questions. The results of Analysis 6 suggest that this difference was not due to differences in the conceptual complexity of argument and adjunct questions. The results of Analysis 7 suggest that the difficulty posed

by adjunct questions is a specific problem with adjunct questions and not a general problem with adjuncts.

3. Why do children acquire questions in this order?

Introduction

The results of the first set of analyses suggest that object questions are acquired before subject questions and the results of the second set of analyses suggest that argument questions are acquired before adjunct questions. What accounts for these two results? Are they both due to a single factor? Two different sorts of factors could account for these results. They could be a reflection of either the syntactic structure of the questions or the learning mechanism of the children.

Structural differences between the different types of questions could cause children to acquire object and argument questions before subject or adjunct questions. According to current GB theory (Chomsky, 1986), object questions differ from subject and adjunct questions in the way the *wh*-trace is governed. GB's Empty Category Principle (ECP) states that all traces must be properly governed. There are, however, two ways a trace can be properly governed. A *wh*-trace can be directly theta-governed by a verb. This is how object *wh*-traces are governed (the simplified structure is shown in i.below). Alternatively, a *wh*-trace can be indirectly or antecedent-governed by a *wh*-word via a chain. This is how subject *wh*-traces and adjunct *wh*-traces are governed (simplified structures are shown in ii. and iii. below).

i.	object	Who_i will [$_{IP}$ Mary [$_{VP}$ meet t_i?]]]
ii.	subject	Who_i [$_{IP}$ t_i [$_{VP}$ meet John?]]]
iii.	adjunct	When_i will [$_{IP}$ Mary [$_{VP}$ meet John t_i?]]]

Theta-government by a verb is more direct and less complicated than antecedent-government by a *wh*-word. (For example, *wh*-words that are directly theta-governed can be extracted more freely than *wh*-words that are antecedent-governed.) It seems plausible, therefore, that questions with theta-governed gaps (i.e., object questions) would be acquired before questions with antecedent-governed gaps (i.e., subject and adjunct questions). I will call this first theory the Government Theory.

Alternatively, children may have some kind of learning procedure which permits them to acquire object questions before subject and adjunct questions. For example, children might ask argument questions first because they notice object gaps more than subject or adjunct gaps. I will call this theory the Visibility Theory.

Let's start with the assumption that children know the argument requirements of verbs. They know, for example, that the verb meet requires that there be someone who meets and someone who is met. When they hear the adjunct question *why did Mary meet John [gap]?*, they can ignore the sentence-initial *why* and its gap and still have a grammatical string (*did Mary meet John?*). When they hear the subject question *who met John?*, they must treat *who* as the subject of the sentence or the string will be ungrammatical (**met John*). They do not, however, have to treat *who* as a question word with a gap. They could treat *who* as some kind of pronoun or proper name.

Object gaps are quite obvious. Upon hearing the object question *who did Mary meet [gap]?*, children might notice that the verb *meet* isn't followed by an object. They might be compelled to search for a suitable object. In their search, they might notice the

who at the front of the sentence and make the connection between the missing object at the end of the sentence and the *who* at the beginning of the sentence. Notice that object gaps are only this blatant in questions with verbs that **require** an object. When children hear object questions with verbs that optionally take objects (e.g., *what did she eat?*), they can assume that these are merely sentences in which the verbs don't take an object. They can ignore the *wh*-word and its trace and still have a grammatical string (*did she eat?*).

Analysis 8. Adult obligatory object questions and child acquisition

The Visibility Theory suggests that parents who ask mostly obligatory object questions will have children who acquire object questions earlier than parents who rarely ask obligatory object questions. The Government Theory does not make this prediction. For each of the 12 children, I determined the percentage of the adults' *who* and *what* object questions that had obligatory objects.

Results. Overall, 98% of adult *who* object questions and 91% of adult *what* object questions had obligatory objects. The percentage of adult obligatory object questions did not correlate with the age of acquisition of either subject or object *who* or *what* questions (all correlation coefficients were between -.13 and +.31).

Summary

In this paper, I determined when 12 children began to ask different types of questions. I found that they began to ask object questions before they began to ask subject questions (Analyses 1-4) and argument questions before adjunct questions (Analyses 5-7). These results could be a reflection of the relative visibility of object traces and adjunct and subject traces (the Visibility Theory). Alternatively, they could be due to differences in how object traces and adjunct and subject traces are governed (the Government Theory). The results of Analysis 8 are more consistent with the Government Theory than the Visibility Theory. In conclusion, the results of all of the transcript analyses are consistent with the hypothesis that children are able to theta-govern traces before they are able to antecedent-govern traces.

---

[1] This research was supported by a National Science Foundation Graduate Fellowship and a MacArthur Foundation M.D/Ph.D Fellowship awarded to the author. I would like to thank Kay Bock, Robin Clark, Brian MacWhinney, Steve Pinker and Ken Wexler for their valuable advice.

[2] Children understood object questions that had *do*-support as well, or better than, they understood subject questions. The only object question that they didn't comprehend as well as subject questions were object questions that had progressive participles (e.g., *what is she kicking*).

[3] I did not include the following 6 types of questions in the subject/object analyses:

1. nonsentential questions (e.g., *what?*)
2. routine questions (e.g., *what's that?*)
3. questions that could be subject or object questions (e.g., *what is a funny shape?*)
4. questions that could be subject or sentential questions (e.g., *what is wrong with Billy?*)
5. questions that questioned a verb (e.g., *what is he doing?*)
6. "*what-for*" questions (e.g., *what did he do that for?*).

[4] All significance levels are for two-tailed tests.

[5] The adult percentages in Analyses 3 and 8 are the percentages for adult questions that were asked prior to the age at which the children acquired the type of question under investigation.

[6] I chose to search for these 3 prepositions because they occur most frequently in the transcripts. It is possible that by using this restricted search I overestimated the age at which the children acquired locative

arguments and adjuncts in declaratives. It seems likely, however, that this search inflates the acquisition age equally for argument and adjunct locatives.

References

Bloom, L. 1973. *One word at a time: The use of single word utterances before syntax.* The Hague: Mouton.

Brown, R. 1973. *A first language: The early stages.* Cambridge, MA: Harvard University Press.

Cairns, H., & Hsu, J. 1978. Who, why, when and how: A developmental study. *Journal of Child Language, 5*, 477-488.

Chomsky, N. 1986. *Barriers.*. Cambridge, MA: MIT Press.

Ervin-Tripp, S. 1970. Discourse agreement: How children answer questions. In J. Hayes (Ed.), *Cognition and the development of language.* New York: Wiley.

Frazier, L., Clifton, C., & Randall, J. 1983. Filling gaps: Decision principles and structure in sentence comprehension. *Cognition, 13*, 187-222.

Gazdar, G. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry, 12*, 155-184.

Gazdar, G., Klein, E., Pullum, G., & Sag, I. 1985. *Generalized phrase structure grammar.* Cambridge, MA: Harvard University Press.

Higginson, R. 1985. Fixing: A demonstration of the child's active role in language acquisition. Unpublished doctoral dissertation. Washington State University.

Labov, W., & Labov, T. 1978.. Learning the syntax of questions. In R. Campbell & P. Smith (Eds.), *Recent advance in the psychology of language: Formal and experimental approaches.* New York: Plenum Press.

MacWhinney, B. n.d. Transcripts of MacWhinney's two children. Available through ChiLDES.

MacWhinney, B., & Snow, C. 1985. The Child Language Data Exchange System. *Journal of Child Language, 12*, 271-296.

Sachs, J. 1983. Talking about there and then: The emergence of displaced reference in parent-child discourse. In K. Nelson (Ed.), *Children's language, Volume 4.* Hillsdale, NJ: Lawrence Erlbaum Press.

Stewart, J., & Sinclair, H. 1975. Comprehension of questions by children between 5 and 9. *International Journal of Psycholinguistics, 3*, 17-26.

Tyack D., & Ingram, D. 1977. Children's production and comprehension of questions. *Journal of Child Language, 4*, 211-228.

Wanner, E., & Maratsos, M. 1978. An ATN approach to comprehension. In M. Halle, J. Bresnan, & G. Miller (eds.), *Linguistic theory and psychological reality.* Cambridge, MA: MIT Press.